arXiv:1508.00683v1 [cs.SY] 4 Aug 2015

# Interval Predictability in Discrete Event Systems

Alban Grastien*

August 5, 2015

**Abstract**

In this paper we study the problem of predictability in partially observable discrete event systems, i.e., the question whether an observer can predict the occurrence of a fault. We extend the definition of predictability to consider the time interval where the fault will occur: the $(i, j)$-predictability does not only specify that the fault will be predicted before it occurs, but also that the predictor will be able to predict that its occurrence will occur in $i$ to $j$ observations from now. We also provide a quadratic algorithm that decides predictability of the system.

**Keywords: Predictability, Discrete Event Systems**

## 1 Motivation

A fault is predictable if its unavoidable occurrence can always be determined in advance. Being able to predict the fault allows the supervisor to step in and take preventive actions, such as reconfiguring the system, replacing damaged components, or shutting the system down.

Predictability has been greatly studied in the last decade (some references are provided in the related work section). To be maximally effective, the prediction should satisfy two criteria: it should be made well in advance, so that the operator has enough time to decide for and perform corrective actions; it should be reasonably precise, so that the repair is not performed too early if that is unnecessary. The first contribution of this paper is the formalisation of these two objectives: we define the notion of $(i, j)$-predictability, a generalisation of the existing notion of predictability that states that faults can always be predicted at least $i$ timesteps in advance and, when this prediction is made, the fault will not occur in more than $j$ timesteps.

We study this definition of predictability and we propose an algorithm that computes all pairs $(i, j)$ for which predictability holds. We show that this algorithm runs in quadratic time. This is an improvement over the existing predictability algorithms that run in $O(n^4)$.

This paper is organised as follows. Next section presents preliminary definitions. Our definition of predictability is presented in Section 3, together with a discussion of its benefits. Our algorithm is given in Section 4. Existing approaches are discussed in Section 5.

## 2 Preliminaries

### 2.1 Discrete Event Systems

This work is applicable to finite discrete event systems (DES) [CL99]. The system is modeled as a DES and is assumed fixed for this paper. A (finite) DES is a model for dynamic systems where the state space is discrete (and finite) and is modeled as a finite state machine.

A (partially observable) *finite state machine* (FSM) is a tuple $A = \langle Q, \Sigma, T, q_{\mathrm{I}}, \Sigma_{\mathrm{o}} \rangle$ where $Q$ is a finite set of states, $\Sigma$ is a finite set of events, $T \subseteq Q \times \Sigma \times Q$ is a finite set of transitions, $q_{\mathrm{I}} \in Q$ is the initial state, and $\Sigma_{\mathrm{o}} \subseteq \Sigma$ is a finite set of observable events.

*A. Grastien is with NICTA, Australia, and the Australian National University.

To simplify notations, it is assumed that the FSM is deterministic, i.e., there is only one initial state and there are no two transitions originating from the same state and labeled with the same event:

$$\{\langle q, e, q'_1\rangle, \langle q, e, q'_2\rangle\} \subseteq T \;\Rightarrow\; q'_1 = q'_2.$$

This assumption is not restrictive as any non-deterministic FSM can be turned into a deterministic FSM that is equivalent from a predictive/monitoring perspective, by adding a number of states and transitions smaller than the original number of transitions and without affecting the overall complexity of the algorithm. Furthermore the algorithms presented later apply to non-deterministic FSM as well. The assumption of determinism is however convenient because there a one-to-one mapping between a path and a trace (defined below).

A *path* $\rho$ is a double sequence of states and events $q_0 \xrightarrow{e_1} \ldots \xrightarrow{e_k} q_k$ such that $\forall i \in \{1, \ldots, k\},\ \langle q_{i-1}, e_i, q_i\rangle \in T$. The label $u$, called the *trace*, of the path is the sequence of events $u = e_1 \ldots e_k$. That there exists a path labeled by $u$ from $q_0$ to $q_k$ is denoted $q_0 \xrightarrow{u} q_k$; the state $q_k$ reached from $q$ through $u$ is denoted $q \xrightarrow{u}$ and the fact that it exists is written $(q \xrightarrow{u}) \in Q$.

The definition of a path is extended to infinite paths $q_0 \xrightarrow{e_1} q_1 \xrightarrow{e_2} \ldots$ such that for all $i \geq 0$, $q_0 \xrightarrow{e_1} \ldots \xrightarrow{e_i} q_i$ is a path. It is assumed that the system is live, i.e., that for any state $q \in Q$, there exists an outgoing transition: $\forall q \in Q,\ \exists e \in \Sigma.\ \exists q' \in Q.\ \langle q, e, q'\rangle \in T$. Infinite traces are denoted $w$ and finite ones $u$. The *prefix* relation is denoted $u \sqsubseteq v$ where $v$ may be finite or infinite. We extend the notation $(q \xrightarrow{w}) \in Q$ to infinite traces, with the meaning $\forall u \sqsubseteq w.\ (q \xrightarrow{u}) \in Q$.

The system starts in state $q_0 = q_\mathrm{I}$ and takes an infinite path. The language $\mathcal{L} = \{w \in \Sigma^\omega \mid (q_\mathrm{I} \xrightarrow{w}) \in Q\}$ is defined as the set of infinite words over $\Sigma$ that label an infinite path on the FSM starting from the initial state.

Given a finite word $u \in \Sigma^*$, the *observation* of $u$ is the traditional projection of $u$ on the set of observable events:

$$\mathrm{obs}(u) = \begin{cases} \varepsilon & \text{if } u = \varepsilon, \\ \mathrm{obs}(u') & \text{if } u = eu' \text{ and } e \in \Sigma \setminus \Sigma_\mathrm{o}, \\ e\ \mathrm{obs}(u') & \text{if } u = eu' \text{ and } e \in \Sigma_\mathrm{o} \end{cases}$$

where $\varepsilon$ is the empty sequence. As usual it is assumed that any infinite trace generates infinitely many observations.

### 2.2 Faults

The system can be subject to faults, i.e., types of behaviour that we wish to prevent. Faults can be defined as a single event or as a subtle pattern of events [JMPC06]. These two definitions are however very similar: the important notion here is that it can also be modeled as the property of the current (possibly augmented) state of the system (normal state vs. faulty state). A set $F \subseteq Q$ of states will represent the faulty states: a path is faulty if it reaches a faulty state ($\exists i.\ q_i \in F$). The faulty aspect of a trace $u$ will therefore be represented by $(q_\mathrm{I} \xrightarrow{u}) \in F$. Notice that, by definition, any transition from a faulty state leads to a faulty state:

$$\langle q, e, q'\rangle \in T \;\wedge\; q \in F \Rightarrow q' \in F.$$

It is assumed that the initial state is not faulty. The set of infinite faulty traces is represented by language $\mathcal{L}_F \subset \mathcal{L}$, which is formally defined as the set of traces whose path from $q_\mathrm{I}$ is faulty.

## 3 $(i, j)$-Predictability

### 3.1 Predictability

Fault prediction is the problem of deciding whether an operator should be warned that a fault is bound to occur. We want to give guarantees about the prediction of the fault. This guarantee is expressed by a tuple $(i, j)$ where $i$ (resp. $j$) is a lower bound (resp. upper bound) of the fault occurrence.

In the following a *time interval* is a pair of elements $(x, y)$ from $\mathbf{N} \cup \{\infty\}$ (the natural numbers including zero and infinity) so that $x \leq y$. We define the operator $\ominus$ so that $(x, y) \ominus 1 = (x \ominus 1, y \ominus 1)$ where $\ell \ominus 1 = \ell$ if $\ell \in \{0, \infty\}$ and $\ell \ominus 1 = \ell - 1$ otherwise. A time interval $(x, y)$ can be interpreted as the set of numbers between $x$ and $y$. Under this interpretation the relation $(x, y) \subseteq (x', y')$ is equivalent to $x' \leq x \leq y \leq y'$; and $(x, y) \cup (x', y') = (\min(x, x'), \max(y, y'))$.[1]

A predictor is a machine $P$ that, given a sequence $o$ of observations, returns a time interval $(x, y) = P(o)$, meaning that any trace that matches this sequence will not become faulty before $x$ more observations are collected (if $x = 0$, the fault may already have occurred) but will definitely be faulty before $y$ more observations are (or returns $y = \infty$ if the fault is not predicted—it may never occur). In the coming definition, notice that, while this is not explicitely stated, if $u$ and $u'$ are two different traces that generate the same observations ($\text{obs}(u) = \text{obs}(u')$) then the predictor should obviously give the same prediction: $P(\text{obs}(u)) = P(\text{obs}(u'))$. Hence the predictor has to be conservative so as to satisfy the two constraints given in the definition for all relevant traces. In other words, there are two types of uncertainty: uncertainty about what happened until now (we only know that the behaviour generated the sequence $o$ but the actual behaviour is unknown); uncertainty about what will happen from now.

**Definition 1** *A* predictor *is a machine $P$ that takes a sequence of observations and that returns a time interval with the following property: $\forall w \in \mathcal{L}$. $\forall u_1, u_2$ such that $u_1 \sqsubseteq u_2 \sqsubseteq w$, let $(x, y) = P(\text{obs}(u_1))$, then*

- $|\text{obs}(u_2)| - |\text{obs}(u_1)| < x \Rightarrow (q_{\text{I}} \xrightarrow{u_2}) \notin F$ *and*
- $|\text{obs}(u_2)| - |\text{obs}(u_1)| \geq y \Rightarrow (q_{\text{I}} \xrightarrow{u_2}) \in F$.

An $(i, j)$-predictor has the added requirement that, before a fault occurs, a prediction should be made about the fault occurrence that is tighter than, or as tight as, $(i, j)$.

**Definition 2** *A predictor $P$ is an* $(i, j)$-predictor for a given trace $w \in \mathcal{L}_F$ *if*

$$\exists u \sqsubseteq w.\ P(\text{obs}(u)) \subseteq (i, j).$$

*A predictor is an* $(i, j)$-predictor *if it is an $(i, j)$-predictor for every trace $w \in \mathcal{L}_F$.*

$(i, j)$-predictability is then the property that an $(i, j)$-predictor exists. We also define *$i$-predictability*, the property that the fault occurrence can be predicted at least $i$ observations before it occurs; and *predictability*, the property that the fault can be predicted before it occurs.

**Definition 3** *A system is* $(i, j)$-predictable *if there exists an $(i, j)$-predictor for it. It is* $i$-predictable *if it is $(i, j)$-predictable for some $j \in \mathbf{N}$. It is* predictable *if it is $i$-predictable for some $i \in \mathbf{N} \setminus \{0\}$.*

Notice that the condition $j \in \mathbf{N}$ (i.e., $j \neq \infty$) is necessary because forbidding the upper bound of $P(o)$ to be $\infty$ forces the predictor to predict the fault before its occurrence (i.e., the predictor asserts that the fault will definitely occur). Similarly we forbid $i = 0$ because we want the fault to be predicted in a state where it has not occurred yet.

| Observation pattern | Prediction |
|---|---|
| No $d$ | $[2, \infty]$ |
| Last observed event is $d$ | $[1, 2]$ |
| Second last observed event is $d$ | $[0, 1]$ |
| Contains $d$ followed by two or more observed events | $[0, 0]$ |

Table 1: A $(1, 2)$-predictor for the system of Figure 1.

[1] Notice that $(x, y) \cup (x', y')$ may contain elements that are neither in $(x, y)$ nor in $(x', y')$.

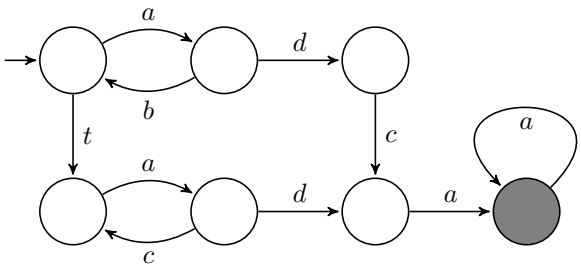



Figure 1: Example of a system; $t$ is the only unobservable event.

These definitions are illustrated with the example of Figure 1. The faulty states are represented with grey filling. Table 1 presents one predictor. For instance the first pattern of the predictor specifies that if the sequence of observations does not contain the event $c$ then the prediction is $(2, \infty)$, i.e., there will be at least two observations before the fault occurs, and it may never occur. The second pattern specifies that if the last event of the sequence of observations is $d$ then the prediction is $(1, 2)$, meaning that a faulty state will be reached after one or two more observations are received. Similarly for the third pattern: the prediction is $(0, 1)$, i.e., it may already have occurred or it will when the next observation has been received. Finally the last pattern indicates a situation where the fault definitely occurred.

We illustrate that the machine in Table 1 (denoted $P$ here) indeed presents a predictor on a few selected examples. We first assume a trace $u_1 = aba$ with prediction $P(\text{obs}(u_1)) = (2, \infty)$. Consider its continuation $u_2 = u_1 b$; then the length difference between $\text{obs}(u_2)$ and $\text{obs}(u_1)$ is 1, which is less than 2; therefore $u_2$ has to satisfy $(q_\text{I} \xrightarrow{u_2}) \notin F$, which it does. Consider instead $u_2 = u_1 dc$; the length difference is this time 2, which means that none of the constraints in Definition 1 applies. Predictor $P$ is not claimed to be "optimal" (where the precise definition of optimality is presented later); nevertheless one might claim that a prediction of $(2, \infty)$ is not very precise given that any continuation of $u_1$ requires three observable events to reach a faulty state ($dca$ is the shortest). Notice however that $P$ does not know that the system trace is $u_1$: it only knows the sequence of observations generated by $u_1$, i.e., $aba$, which is identical to the sequence generated by $u_1' = abta$; this trace $u_1'$ can reach a faulty state in just two observable steps ($da$), which forces the lower bound of $P(\text{obs}(u_1))$ to be at most 2.

Assume now $u_1 = abad$ with prediction $(1, 2)$. Consider the non-faulty trace $u_2 = u_1 c$; the length difference is 1, which means that none of the constraints in Definition 1 applies. Consider instead the faulty trace $u_2 = u_1 ca$; the length difference is 2, which is greater or equals to the upper bound of the prediction; therefore $u_2$ has to satisfy $(q_\text{I} \xrightarrow{u_2}) \in F$, which it does.

As we can see any faulty trace has to include $d$, which means that the flow of observations generated by a faulty trace will eventually be associated with the prediction $(1, 2)$. Therefore the system is $(1, 2)$-predictable. We can however show that the system is not $(2, 2)$-predictable. Indeed consider the infinite faulty trace $w = adca^\omega$ where the exponent $^\omega$ indicates an infinite repetition of $a$. For $w$ to be $(2, 2)$-predictable, we need to exhibit one of its prefix $u_1$ such that one can predict $P'(\text{obs}(u_1)) \subseteq (2, 2)$ (here $P'(\text{obs}(u_1))$ should exactly equal $(2, 2)$). Assume that such a prefix and such a predictor exist. Following Definition 1, consider a continuation $u_2$ of $u_1$ that generates one more observation; because $|\text{obs}(u_2)| - |\text{obs}(u_1)| = 1$, $u_2$ should not lead to a faulty state. Therefore $u_1$ has to belong to the set $\{\varepsilon, a, ad\}$. Similarly however, if $u_2$ is chosen such that its observable length is exactly two more than that of $u_1$, then $u_2$ has to lead to a faulty state. Therefore $u_1 = ad$ and $P'(\text{obs}(u_1)) = P'(ad) = (2, 2)$. Consider however the trace $u_1' = tad$ and its continuation $u_2' = u_1' a$. Clearly $P'(\text{obs}(u_1')) = P'(ad) = (2, 2)$. According to Definition 1 since $|\text{obs}(u_2')| - |\text{obs}(u_1')| = 1 < 2$ $u_2'$ should not lead to a faulty state. It does however, which shows that no prefix $u_1$ of $w$ satisfies $P'(u_1) \subseteq (2, 2)$ for some predictor $P'$.

### 3.2 Discussion

Predictors can be used to stop or rectify the system before it produces a faulty behaviour. Being able to predict a fault well in advance helps getting prepared for intervention; this is represented by the $i$ parameter (which should be maximised). Being able to predict the time when the fault is likely to happen prevents hasty corrections; this is represented by the difference $(j - i)$ (which should be minimised). There is an implicit assumption here that the number of observations is indicative of time: for instance the system generates one observation per minute. This is particularly relevant to hybrid systems modeled as DES [VTPS15].

Ideally the system should be $(i, j)$-predictable with a large $i$ value and a small $(j - i)$ value.

We illustrate the definition of predictability by considering the example of the potentially critical subsystem of an aircraft. This example is, of course, very limited. For such a system it is important to predict faults well in advance in order to take preventive measures (e.g., modify the flight path in order to stay near to an aerodrome). On the other hand it is also important to provide a precise prediction as emergency landings are expansive.

In order to provide an early prediction we might want the system to be at least 30-predictable. At that stage however, we do not need a precise prediction: a $(30, 10\,000+)$-predictability is still acceptable. For the second requirement however, we want to be able to predict the fault quite accurately, for instance $(15, 240)$-predictability which suggests that the fault will occur in the next four hours and that an unscheduled landing is now necessary. So, interestingly, this example requires two different predictability properties.

## 4 Solving Interval Predictability Problems

This section shows how to verify the predictive level of a given system.

### 4.1 Predictive levels

We first show that, while the definition of predictability involves two parameters, the dimension of predictability is actually much smaller.

**Lemma 1** *A system that is $(i, j)$-predictable is also*

1. *$(i, (j + 1))$-predictable (if $j \neq \infty$) and*
2. *$((i - 1), (j \ominus 1))$-predictable (if $i \geq 2$).*

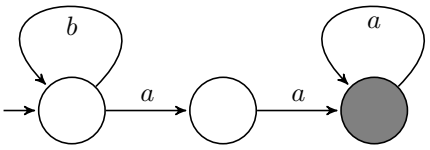
**Proof** That $(i, j)$-predictability entails $(i, j+1)$-predictability is trivial from Definition 2: an $(i, j)$-predictor is also an $(i, j + 1)$-predictor since the constraint on the prediction is strictly weaker.

Assume that the system is $(i, j)$-predictable with $i \geq 2$, i.e., there exists an $(i, j)$-predictor $P$. Then define $P'$ such that

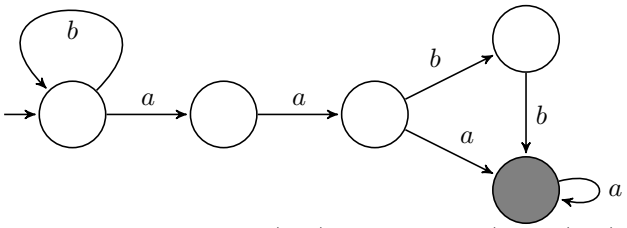
- $P'(\varepsilon) = P(\varepsilon)$ and
- $P'(oe) = P(o) \ominus 1$.

It is easy to show that $P'$ is a predictor (if the prediction $P(o)$ was correct, then the prediction $P'(oe)$ is correct). Furthermore it is easy to prove that $P'$ is an $(i-1), (j-1)$-predictor: if $P(\text{obs}(u)) \subseteq (i, j)$ for some prefix $u$ of $w$, then for the prefix $ue$, $P'(ue) = P(u) \ominus 1 \subseteq (i - 1, j - 1)$ (or $(i - 1, \infty)$ if $j = \infty$). □

Lemma 1 shows that some levels of predictability are strictly weaker than others. There are however levels of predictability that are mutually incomparable. Consider the examples of Figure 2. Clearly the system of Figure 2a is $(1, 1)$-predictable because a fault is always preceded by two $a$s and the occurrence of the first $a$ implies that the fault will be reached after the next observation; on the other hand it is not $(2, 3)$-predictable because when the fault becomes unavoidable (i.e., it will occur after less than 3 observations) then the fault can (and, actually, will) occur after less than 2 observations. The system of Figure 2b is $(2, 3)$-predictable because the fault is always preceded by $aaa$ or $aabb$ and because observing a first $a$ implies that the fault is unavoidable; on the other hand, it is not $(1, 1)$-predictable because, after observing $aa$, it is not possible to decide whether the fault will occur immediately or after two observations.




a. Example of a $(1,1)$-predictable set of faults which is not $(2,3)$-predictable.




b. Example of a $(2,3)$-predictable (and $(1,2)$-predictable) set of faults which is not $(1,1)$-predictable.

Figure 2: Illustrating that some predictive levels are not comparable.

## 4.2 Characterisation of Predictability

In order to determine whether a system is predictable we define notions of distance between a system state and a fault.

**Definition 4** *The* minimal distance *between $q$ and the set $F$ of states denoted $dmin_F(q)$, is the minimum number of observations before reaching $F$ from $q$*

$$dmin_F(q) = \min_{(q \xrightarrow{u}) \in F} |\mathrm{obs}(u)|$$

*and $\infty$ if there is no such $u$. The* maximal distance *between $q$ and a set of states $F$, denoted $dmax_F(q)$, is the maximum number of observations before reaching $F$ from $q$*

$$dmax_F(q) = \max_{(q \xrightarrow{u}) \in Q \setminus F} |\mathrm{obs}(u)| + 1,$$

*$\infty$ if there is no bound to $|\mathrm{obs}(u)|$, and $0$ if $q \in F$ (i.e., there is no such $u$).*

Notice that these distances are bounded by the number $|Q|$ of states when they are different from $\infty$. Indeed, if $dmin_F(q) \geq |Q|$ the corresponding trace includes a cycle, and a smaller trace therefore exists (by cutting the cycle). Similarly if $dmax_F(q) \geq |Q|$ the corresponding trace includes a cycle, and a longer trace exists (where the cycle can be taken once more).

The minimal and maximal distances give us a first estimate of the time interval before fault. To simplify notations we write $distances_F(q)$ to denote the time interval $(dmin_F(q), dmax_F(q))$.

**Lemma 2** *For all trace $w \in \mathcal{L}$ and all prefix $u \sqsubseteq w$, if $P$ is a predictor then*

$$distances_F(q_{\mathrm{I}} \xrightarrow{u}) \ \subseteq \ P(\mathrm{obs}(u)).$$

**Proof** Let $(i,j) = distances_F(q_{\mathrm{I}} \xrightarrow{u})$ be the time interval of the state $(q_{\mathrm{I}} \xrightarrow{u})$ and let $(x,y) = P(\mathrm{obs}(u))$ be the prediction of $\mathrm{obs}(u)$.

By definition of the minimal distance $i$, there exists a trace $u'$ such that $u \sqsubseteq u' \sqsubseteq w$, $|\mathrm{obs}(u')| - |\mathrm{obs}(u)| = i$, and $(q_{\mathrm{I}} \xrightarrow{u'}) \in F$. Therefore $i < x$ would contradict the first condition in the definition of a predictor (Def. 1).

Furthermore if $j > 0$, then by definition of the maximal distance $j$, there exists a trace $u'$ such that $u \sqsubseteq u' \sqsubseteq w$, $|\mathrm{obs}(u')| - |\mathrm{obs}(u)| = j - 1$, and $(q_{\mathrm{I}} \xrightarrow{u'}) \in Q \setminus F$. Therefore $j - 1 \geq y$ would contradict the second condition in the definition of a predictor (Def. 1).

If $j = 0$ then $i = 0$ and $y \geq x \geq i = j$ implies $y \geq j$. □

This result can be generalised to the collection of states that an observer can assume the system to be in (the *belief state*). Formally the belief state $\mathcal{B}(o)$ is the set of states that the system can be in if the sequence $o$ of observations has been observed:

$$\mathcal{B}(o) = \{q \in Q \mid \exists u.\ q_{\mathrm{I}} \xrightarrow{u} q \ \wedge \ \mathrm{obs}(u) = o\}.$$

**Corollary 3** *For all predictor $P$, for all sequence $o$ of observations*

$$\left( \bigcup_{q \in \mathcal{B}(o)} \mathit{distances}_F(q) \right) \subseteq P(o).$$

**Proof** If $q \in \mathcal{B}(o)$ is an element of the belief state then, by definition of the belief state, there exists a trace $u$ such that $q_{\mathrm{I}} \xrightarrow{u} q$ and $o = \mathrm{obs}(u)$. From Lemma 2, $\mathit{distances}_F(q) \subseteq P(o)$, which also applies to the union of these elements. □

Actually it is possible to characterise the "optimal" predictor in terms of distances. Let $P$ and $P'$ be two predictors. We say that $P$ is stronger than $P'$, denoted $P \succeq P'$, iff $P(o) \subseteq P'(o)$ for all $o$.[2] We denote $P^\star$ the *optimal predictor*: $P^\star = \max_{\succeq}\{P \mid P \textit{ is a predictor}\}$. It should be clear that the optimal predictor is well-defined and unique.

**Lemma 4** *The optimal predictor $P^\star$ is exactly the predictor that satisfies $P^\star(o) = \left(\bigcup_{q \in \mathcal{B}(o)} \mathit{distances}_F(q)\right)$ for all sequence $o$ of observations.*

**Proof** Let $(i, j) = \left(\bigcup_{q \in \mathcal{B}(o)} \mathit{distances}_F(q)\right)$ and $(x, y) = P^\star(o)$. From Corollary 3 we already know that $(i, j) \subseteq (x, y)$. We only need to prove that $P(o) = (i, j)$ is a correct prediction.

Following Definition 1 let $w \in \mathcal{L}$ be an infinite trace and let $u_1, u_2$ be two finite traces such that $u_1 \sqsubseteq u_2 \sqsubseteq w$ and $\mathrm{obs}(u_1) = o$. Let us call $q_1$ the state reached by $u_1$ and $q_2$ the state reached by $u_2$: $q_{\mathrm{I}} \xrightarrow{u_\ell} q_\ell$. By definition of the belief state, $q_1 \in \mathcal{B}(o)$. To prove that $P(o)$ is a correct prediction we need to prove that the two conditions of Definition 1 are satisfied.

Assume that $q_2 \notin F$; we shall prove that the premise of the second condition in Definition 1 is not satisfied. By definition of the maximal distance of $q_1$: $\mathit{dmax}_F(q_1) > |\mathrm{obs}(u_2)| - |\mathrm{obs}(u_1)|$. Since we know $j \geq \mathit{dmax}_F(q_1)$, it clearly holds that $|\mathrm{obs}(u_2)| - |\mathrm{obs}(u_1)| < j$.

Assume instead that $q_2 \in F$; we shall prove this time that the premise of the first condition is not satisfied. By definition of the minimal distance of $q_1$: $\mathit{dmin}_F(q_1) \leq |\mathrm{obs}(u_2)| - |\mathrm{obs}(u_1)|$. Since we know $i \leq \mathit{dmin}_F(q_1)$, it clearly holds that $|\mathrm{obs}(u_2)| - |\mathrm{obs}(u_1)| \geq i$. □

As it turns out $P^\star(o)$ equals the union of exactly two intervals.

**Lemma 5** *For all sequence $o$ of observations such that $\mathcal{B}(o) \neq \emptyset$, there exists a pair of states $\{q_1, q_2\} \subseteq \mathcal{B}(o)$ such that $P^\star(o) = \mathit{distances}_F(q_1) \cup \mathit{distances}_F(q_2)$.*

**Proof** From Lemma 4 $P^\star(o)$ is the union of a finite collection of intervals. Because this set is finite, there is an interval, say $\mathit{distances}_F(q_1)$, whose lower bound is minimal; similarly there is an interval, say $\mathit{distances}_F(q_2)$, whose upper bound is maximal. Therefore $P^\star(o) = \mathit{distances}_F(q_1) \cup \mathit{distances}_F(q_2)$. □

The optimal predictor exhibits some very interesting properties.

**Lemma 6** *For all sequence $o$ of observations,*

$$P^\star(oe) \subseteq P^\star(o) \ominus 1.$$

**Proof** Let $u_1 \sqsubseteq u_2$ be two finite traces such that $|\mathrm{obs}(u_2)| = |\mathrm{obs}(u_1)| + 1$. Then by definition $\mathit{dmin}_F(q_{\mathrm{I}} \xrightarrow{u_1}) \geq \mathit{dmin}_F(q_{\mathrm{I}} \xrightarrow{u_2}) + 1$ (unless $\mathit{dmin}_F(q_{\mathrm{I}} \xrightarrow{u_1}) = 0$). Similarly $\mathit{dmax}_F(q_{\mathrm{I}} \xrightarrow{u_1}) \leq \mathit{dmax}_F(q_{\mathrm{I}} \xrightarrow{u_2}) + 1$ (unless $\mathit{dmax}_F(q_{\mathrm{I}} \xrightarrow{u_1}) = \infty$).

Therefore $\mathit{distances}_F(q_{\mathrm{I}} \xrightarrow{u_1}) \subseteq \mathit{distances}_F(q_{\mathrm{I}} \xrightarrow{u_2}) \ominus 1$.

For each state $q_2 \in \mathcal{B}(oe)$, there exists a state in $q_1 \in \mathcal{B}(o)$ such that two such traces $u_1 \sqsubseteq u_2$ lead respectivement to $q_1$ and $q_2$ (but notice that for some $q_1$, there may be no such $q_2$). Therefore $P^\star(oe) = \bigcup_{q_2 \in \mathcal{B}(oe)} \mathit{distances}_F(q_2) \subseteq \bigcup_{q_1 \in \mathcal{B}(oe)} \mathit{distances}_F(q_1) \ominus\ 1 = P^\star(o) \ominus 1$. □

[2] We assume that $P(o)$ and $P'(o)$ are undefined if $o$ cannot be generated by the system ($\mathcal{B}(o) = \emptyset$).

The optimal predictor can be used to decide predictability. Indeed from Definition 2 any suboptimal predictor enjoys only a (non-necessarily strict) subset of $(i,j)$-predictability qualities of the optimal predictor. This is expressed in the following corollary where non-predictability is proved if $(i,j)$ is a strict subset ($\subset$) of some prediction $P^\star(o)$.

**Corollary 7** *If $dmin_F(q_\mathrm{I}) \geq i$ the system is not $(i,j)$-predictable iff there exists a sequence $o$ of observations such that $(i,j) \subset P^\star(o)$.*

**Proof** We assume $dmin_F(q_\mathrm{I}) \geq i$.

$\Leftarrow$ Assume that there is no sequence $o$ of observations such that $(i,j) \subset P^\star(o)$. Consider a faulty trace $w$. We shall show that $w$ is $(i,j)$-predictable.

Let $u \sqsubseteq w$ be a faulty prefix: $(q_\mathrm{I} \xrightarrow{u}) \in F$. Then by Definition 1 of a predictor, $P^\star(\mathrm{obs}(u)) = (x,y)$ where $x = 0$. Notice also that $P^\star(\mathrm{obs}(\varepsilon)) = (x^\varepsilon, y^\varepsilon)$ where $x^\varepsilon \geq i$. From Lemma 6 we know that adding one observation to a sequence can reduce the lower bound of the interval returned by $P^\star$ only by 1. Therefore, since the lower bound is greater than or equal to $i$ for $\varepsilon$ and down to 0 for $\mathrm{obs}(u)$, there is a prefix $u'$ of $u$ such that $P^\star(u') = (x',y')$ and $x' = i$. But since $(i,j) \not\subset (x',y')$, $(x',y') \subseteq (i,j)$ and the faulty trace $w$ is $(i,j)$-predictable (Def. 2).

$\Rightarrow$ Let $o$ be the sequence of observations such that $(i,j) \subset P^\star(o)$ and let $(x,y)$ be this interval $P^\star(o)$. Notice that $y \geq 1$ since $(i,j)$ is not empty.

Assume $y \neq \infty$. From Lemma 5 and from the definition of the time intervals there exists $u_1 \sqsubseteq u_2 \sqsubseteq w$ and $u_1' \sqsubseteq u_2' \sqsubseteq w'$ such that

- $\{w, w'\} \in \mathcal{L}$,
- $\mathrm{obs}(u_1) = \mathrm{obs}(u_1') = o$,
- $|\mathrm{obs}(u_2)| - |\mathrm{obs}(u_1)| = x$,
- $(q_\mathrm{I} \xrightarrow{u_2}) \in F$,
- $|\mathrm{obs}(u_2')| - |\mathrm{obs}(u_1')| = y - 1$,
- $(q_\mathrm{I} \xrightarrow{u_2'}) \in Q \setminus F$.

We shall prove by contradiction that $w$ is not $(i,j)$-predictable.

Assume that $w$ is $(i,j)$-predictable. Then there exists a prefix $u_3$ of $w$ such that $P^\star(u_3) \subseteq (i,j)$. Because of the first condition of Definition 1, this prefix must be such that $|\mathrm{obs}(u_2)| - |\mathrm{obs}(u_3)| \geq i$, and therefore $|\mathrm{obs}(u_1)| - |\mathrm{obs}(u_3)| \geq 0$. We know that $P^\star(\mathrm{obs}(u_1)) \not\subseteq (i,j)$, therefore $|\mathrm{obs}(u_1)| - |\mathrm{obs}(u_3)| \geq 1$ and $u_3 \sqsubseteq u_1$.

Because $u_1$ and $u_1'$ generate the same sequence of observations, there exists a prefix $u_3'$ of $u_1'$ (and therefore of $u_2'$) that generates the same sequence of observations as $u_3$. Furthermore, we know that $|\mathrm{obs}(u_2')| - |\mathrm{obs}(u_3')| = |\mathrm{obs}(u_2')| - |\mathrm{obs}(u_3)| > |\mathrm{obs}(u_2')| - |\mathrm{obs}(u_1)| = |\mathrm{obs}(u_2')| - |\mathrm{obs}(u_1')| = y - 1$; that is: $|\mathrm{obs}(u_2')| - |\mathrm{obs}(u_3')| \geq y$. According to the second condition of Definition 1, $(q_\mathrm{I} \xrightarrow{u_3'}) \in F$, which contradicts the last item of the six items presented at the beginning of this proof.

The proof under the assumption that $y = \infty$ is very similar. We choose $u_2'$ such that $|\mathrm{obs}(u_2')| - |\mathrm{obs}(u_1')| > |Q| + 2$. This proves that the system is not $(i, |Q|+1)$-predictable. Since we know that a bound bigger than $|Q|$ is equivalent to that of $\infty$, we show that the system is not $(i,\infty)$-predictable. □

Notice that if $dmin_F(q_\mathrm{I}) < i$, then the system is not $(i,j)$-predictable for any $j$ (even $j = \infty$).

Combining Corollary 7 and Lemma 5, we obtain the following theorem.

**Theorem 8** *The system is $(i,j)$-predictable iff $dmin_F(q_\mathrm{I}) \leq i$ and for all sequence $o$ of observations, for all pair of states $(q_1,q_2) \subseteq \mathcal{B}(o)$,*

$$(i,j) \not\subset distances_F(q_1) \cup distances_F(q_2).$$

We write $q_1 \sim q_2$ the relation indicating that the two states $q_1$ and $q_2$ appear together in a belief state. Notice that $\sim$ is not an equivalence relation (it is not transitive).

### 4.3 Algorithms

We now turn to implementation of Theorem 8. The algorithm includes four steps:

1. Compute the minimal distance for each state;
2. Compute the maximal distance for each state;
3. Compute the twin plant which represents the $\sim$ relation;
4. Compute the $(i, j)$-predictability.

All parts of the verification process will be presented here to ensure the paper is self-contained.

Algorithm 1 computes the minimal distance of each state. In this algorithm and the following one, $c(e) = 1$ if $e$ is observable and 0 otherwise. It assumes that all states have infinite distance until it is has been proved that a shorter distance exists. It then sets all faulty states' minimal distance to 0 and updates the minimal distances of all states until convergence is reached. To make sure that the states are explored in the optimal order we use a priority queue $\mathcal{Q}$ that orders its elements by smaller value $dmin_F(q)$; however since $\mathcal{Q}$ only contains elements with two types of distances (the current distance and this distance plus one), the queue can be implemented with two buckets. The complexity of the algorithm is therefore linear in the number $|T|$ of transitions.

**Algorithm 1** Computing the minimal distance.

**Input**: an FSM $\langle Q, \Sigma, T, q_I, \Sigma_o \rangle$, a set of states $F \subseteq Q$
Create a table $dmin_F : Q \to \mathbf{N} \cup \{\infty\}$
**for all** $q \in Q$ **do**
  $dmin_F(q) := \infty$
**end for**
$\mathcal{Q} = \emptyset$
**for all** $q \in F$ **do**
  $dmin_F(q) := 0$
  $\mathcal{Q} := \mathcal{Q} \cup \{q\}$
**end for**
**while** $\mathcal{Q} \neq \emptyset$ **do**
  $q' := \text{pop}(\mathcal{Q})$
  **for all** $\langle q, e, q' \rangle \in T$ **do**
    **if** $dmin_F(q) > dmin_F(q') + c(e)$ **then**
      $dmin_F(q) := dmin_F(q') + c(e)$
      $\mathcal{Q} := \mathcal{Q} \cup \{q\}$
    **end if**
  **end for**
**end while**
**return** $dmin_F$

Algorithm 2 computes the maximal distance of each state. It starts by computing the list of states ($N$) that can stay outside of $F$ forever (those states have infinite maximal distance). It then initialises every state with a maximal distance of 0 and updates the distance whenever it finds a bigger value. This update will eventually terminate (after at most $|Q|$ iterations). The first part of the algorithm requires to iterate twice over all transitions; the second part requires to iterate at most $|Q|$ times over at most all transitions. Therefore the complexity of Algorithm 2 is at most $|Q| \times |T|$.

**Algorithm 2** Computing the maximal distance.

**Input**: an FSM $\langle Q, \Sigma, T, q_{\mathrm{I}}, \Sigma_{\mathrm{o}} \rangle$, a set of states $F \subseteq Q$
Let $N := Q \setminus F$
Create map nsucc : $N \to \mathbf{N}$
$R := \emptyset$
**for all** $q \in N$ **do**
  nsucc$[q] := |\{\langle q, e, q' \rangle \in T\}|$
  **if** nsucc$[q] = 0$ **then**
    $R := R \cup q$
  **end if**
**end for**
**while** $R \neq \emptyset$ **do**
  Let $q' := \mathrm{pop}(R)$
  **for all** $\langle q, e, q' \rangle \in (N \times \Sigma \times \{q'\})$ **do**
    nsucc$[q]$ := nsucc$[q] - 1$
    **if** nsucc$[q] = 0$ **then**
      $R := R \cup q$
    **end if**
  **end for**
**end while**
Create a table $dmax_F : Q \to \mathbf{N} \cup \{\infty\}$
**for all** $q \in Q$ **do**
  **if** $q \in N$ **then**
    $dmax_F(q) := \infty$
  **else**
    $dmax_F(q) := 0$
  **end if**
**end for**
*needsUpdate* := *true*
**while** *needsUpdate* **do**
  *needsUpdate* := *false*
  **for all** $q' \in Q \setminus N$ **do**
    **for all** $\langle q, e, q' \rangle \in T$ **do**
      **if** $dmax_F(q) < dmax_F(q') + c(e)$ **then**
        $dmax_F(q) := dmax_F(q') + c(e)$
        *needsUpdate* := *true*
      **end if**
    **end for**
  **end for**
**end while**
**return** $dmax_F$

The twin plant [JHCK01] is a construction that determines precisely the $\sim$ relation. Notice that, strictly speaking, it is not necessary to build it as a finite state machine: for predictability only the $\sim$ relation matters; not the transitions between the states of the twin plant.

Given an FSM $A = \langle Q, \Sigma, T, q_{\mathrm{I}}, \Sigma_{\mathrm{o}} \rangle$, the *twin plant* is the finite state machine $\langle Q^T, \Sigma^T, T^T, q_{\mathrm{I}}^T, \Sigma_{\mathrm{o}} \rangle$ where

- $Q^T = Q \times Q$,
- $\Sigma^T = \big((\Sigma \setminus \Sigma_{\mathrm{o}}) \times \{1, 2\}\big) \ \cup \ \Sigma_{\mathrm{o}}$,
- $T^T =$

$$\begin{array}{rl} & \{\langle\langle q_1, q_2\rangle, e^T, \langle q_1', q_2'\rangle\rangle \mid \ \langle q_1, e^T, q_1'\rangle \in T \ \wedge \ \langle q_2, e^T, q_2'\rangle \in T \ \wedge \ e^T \in \Sigma_{\mathrm{o}} \ \} \\ \cup & \{\langle\langle q_1, q_2\rangle, e^T, \langle q_1', q_2'\rangle\rangle \mid \ \langle q_1, e^T, q_1'\rangle \in T \ \wedge \ q_2 = q_2' \ \wedge \ e^T \in \Sigma \setminus \Sigma_{\mathrm{o}} \} \\ \cup & \{\langle\langle q_1, q_2\rangle, e^T, \langle q_1', q_2'\rangle\rangle \mid \ q_1 = q_1' \ \wedge \ \langle q_2, e^T, q_2'\rangle \in T \ \wedge \ e^T \in \Sigma \setminus \Sigma_{\mathrm{o}} \}. \end{array}$$

- $q_{\mathrm{I}}^T = \langle q_{\mathrm{I}}, q_{\mathrm{I}} \rangle$.

It is well-known that the state $\langle q_1, q_2 \rangle$ of $Q^T$ is reachable from $q_{\mathrm{I}}^T$ iff $q_1 \sim q_2$ (Lemma 10, [GL09] where the twin plant is called *verifier*). The twin plant can therefore be used to verify the $(i, j)$-predictability.

The procedure for computing the $(i, j)$-predictability is given in Algorithm 3. It generates an array $p$ such that for all $i$, the system is $(i, p[i])$-predictable and non-$(i, p[i] - 1)$-predictable. The algorithm first initialises $p[i]$ to $i$. It then iterates over all the states of the twin plant and updates the table $p$. According to Theorem 8 the result of Algorithm 3 is the list of $(i, p[i])$-predictabilities that the system enjoys.

**Algorithm 3** Algorithm to compute $(i, j)$-predictability

1: **input** an FSM $A = \langle Q, \Sigma, T, q_{\mathrm{I}}, \Sigma_{\mathrm{o}} \rangle$, the list of minimal and maximal distances $dmin_F$ and $dmax_F$, the twin plant $\langle Q^T, \Sigma^T, T^T, q_{\mathrm{I}}^T, \Sigma_{\mathrm{o}} \rangle$
2: Create an array of integers $p : \mathbf{N}^{|\min F(q_{\mathrm{I}})|}$
3: **for all** $i \in \{1, \ldots, |Q|\}$ **do**
4: $\quad p[i] := i$
5: **end for**
6: **for all** $\langle q, q' \rangle \in \mathrm{Reachable}(Q^T)$ **do**
7: $\quad i := \min(dmin_F(q), dmin_F(q'))$
8: $\quad j := \max(dmax_F(q), dmax_F(q'))$
9: $\quad p[i] := \max(p[i], j)$
10: **end for**
11: **return** $p$

We claim that the algorithm presented here is quadratic in the number of states and transitions of the system. It is easy to see that computing the distances is at most quadratic for both types of distances, and that the resulting structure has linear size with constant time access. The size of the twin plant is quadratic in the size of the original system (it includes at most $|Q|^2$ states and $(2 \times |T| \times |Q|) + |T|^2$ transitions)—and that is assuming a non-deterministic model. Finally the fourth step requires iterating over the quadratic number of states in the twin plant.

Our definitions of predictability and $i$-predictability match those of Jéron et al. [JMGL08] and the proposed algorithm can therefore be used to verify these properties. It is also possible to simplify it by focussing on the $i$ parameter.

As a last result, consider a *fully observable system*, i.e., a system in which $\mathrm{obs}(u) = u$. Then, at any time, the state of the system can be deduced from the sequence of observations; but notice that how the system will evolve remains unknown. Then the relation $\sim$ is equivalent to identity: $q \sim q'$ iff $q = q'$. Consequently, after the distances of each state have been computed, the predictability can be computed in linear time.

### 4.4 Building the Optimal Predictor

Lemma 4 gives us a procedure for computing the optimal predictor. Similarly to diagnosis and its diagnoser [SSL+95] it is possible to compute a deterministic FSM that represents how the belief state evolves as more observations are gathered.

Formally the *optimal predictor* is a finite state machine $\langle Q^\star, \Sigma^\star, T^\star, {q_\mathrm{I}}^\star \rangle$ where

- $Q^\star = \{q^\star \mid q^\star \subseteq Q\}$,
- $\Sigma^\star = \Sigma_\mathrm{o}$,
- $T^\star \subseteq Q^\star \times {\Sigma_\mathrm{o}}^\star \times Q^\star$ is defined below, and
- ${q_\mathrm{I}}^\star = \{q_\mathrm{I}\}$.

For every state ${q_1}^\star \in Q^\star$ of the optimal predictor and every event $e \in \Sigma^\star$, there is exactly one state ${q_2}^\star$ such that ${q_1}^\star \xrightarrow{e} {q_2}^\star$ is a transition of the optimal predictor. The state ${q_2}^\star \subseteq Q$ is defined as the set of states of the system that can be reached from a state of ${q_1}^\star$ through a path that generates only one observation:

$$ {q_2}^\star = \{q_2 \in Q \mid \exists q_1 \in {q_1}^\star.\ \exists u.\ (q_1 \xrightarrow{u} q_2) \wedge (|\mathrm{obs}(u)| = 1)\}. $$

Given a sequence $o$ of observations the predictor follows the single path labeled by $o$ on the predictor and reaches the state $q^\star(o)$ (i.e., the state $q^\star(o)$ such that ${q_\mathrm{I}}^\star \xrightarrow{o} q^\star(o)$). The prediction is then $\bigcup_{q \in q^\star(o)} \mathit{distances}_F(q)$.[3] Adding a single observation $e$ to $o$, the new prediction can be easily computed by getting the state $q^\star(oe)$ that satisfies $q^\star(o) \xrightarrow{e} q^\star(oe)$. Assuming the optimal FSM and the interval associated with each state of the predictor are precomputed, the optimal prediction of a sequence of observations is linear in the size of this sequence and the incremental optimal prediction is constant time. Notice however that, as is the case with the diagnoser [Rin07], the optimal predictor is exponentially large in the number of states of the system.

## 5 Related Work

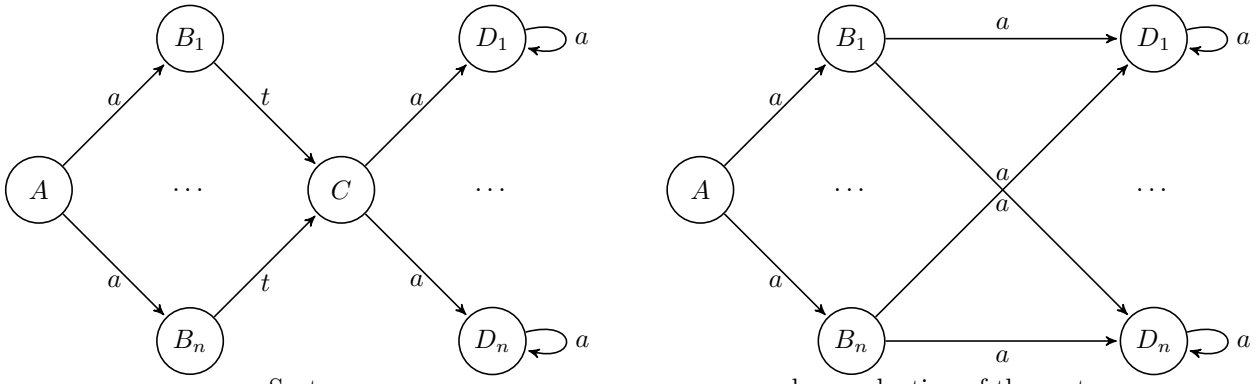
Predictability as presented in this paper was introduced by Genc and Lafortune [GL06]. Their approach was however only Boolean: they addressed the question "can the fault be predicted before it occurs?" They presented an exponential space algorithm using a structure similar to our optimal predictor. They also announced the existence of a polytime algorithm, similar to the twin plant used for diagnosability and formally presented in an extension of their work [GL09].

Together with Jéron and Marchand, they proposed an additional improvement to lower the complexity down to quadratic [JMGL08]. We claim here that their algorithm is not quite quadratic (we discuss this question at the end of this section). Their approach is very similar to the approach presented in the previous section: They construct a twin plant and verify predictability by checking whether there exists a pair $q_1 \sim q_2$ such that $\mathit{dmin}_F(q_1) = 0$ and $\mathit{dmax}_F(q_2) = \infty$.

Brandán Briones and Madalinski presented the notions of *lb*-predictability and *ub*-predictability [BM11]. *ub*-predictability is similar to our definition of $i$-predictability meaning that the fault is predicted at least $i$ observations before the fault occurs. *lb*-predictability is the equivalent of our property of $(1, j)$-predictability, meaning that it is possible to predict the fault occurrence before it occurs but when at most $j$ observations are still possible before the fault (in other words, the fault prediction is not too early).

While this is a minor issue, we provide an example and a comprehensive discussion that illustrate the complexity error from Jéron et al. [JMGL08]. Consider the example of Figure 3a. This DES includes $2n + 2$ states and $4n$ transitions. The single observable event is $a$ and the single unobservable event is $t$ (this example does not feature any faulty event). The twin plant then consists in $2n^2 + 2$ states and $4n^2$

[3] If the model is correct, then the state $q^\star(o) = \{\}$ should never be reached and the union is therefore well-defined.




a. System. b. $\varepsilon$-reduction of the system.

Figure 3: DES (a) and its $\varepsilon$-reduction (b).

| Type of states | Number of states |
|---|---|
| $\langle A, A\rangle$ | 1 |
| $\langle B_i, B_j\rangle$ | $n^2$ |
| $\langle C, C\rangle$ | 1 |
| $\langle D_i, D_j\rangle$ | $n^2$ |
| Total: | $2n^2+2$ |

| Type of transitions | Number of transition |
|---|---|
| $\langle A, A\rangle \to \langle B_i, B_j\rangle$ | $n^2$ |
| $\langle B_i, B_j\rangle \to \langle C, C\rangle$ | $n^2$ |
| $\langle C, C\rangle \to \langle D_i, D_j\rangle$ | $n^2$ |
| $\langle D_i, D_j\rangle \to \langle D_i, D_j\rangle$ | $n^2$ |
| Total: | $4n^2$ |

Table 2: Size of the twin plant for the DES in Figure 3a.

| Type of states | Number of states |
|---|---|
| $\langle A, A\rangle$ | 1 |
| $\langle B_i, B_j\rangle$ | $n^2$ |
| $\langle D_i, D_j\rangle$ | $n^2$ |
| Total: | $2n^2+1$ |

| Type of transitions | Number of transition |
|---|---|
| $\langle A, A\rangle \to \langle B_i, B_j\rangle$ | $n^2$ |
| $\langle B_i, B_j\rangle \to \langle D_k, D_\ell\rangle$ | $n^4$ |
| $\langle D_i, D_j\rangle \to \langle D_i, D_j\rangle$ | $n^2$ |
| Total: | $n^4+2n^2$ |

Table 3: Size of the twin plant for the $\varepsilon$-reduced DES in Figure 3b.

transitions (details in Table 2). The $\varepsilon$-reduction, presented on Figure 3b, contains one state fewer than the original DES but $n^2+2n$ transitions. As a consequence, the number of states in the twin plant reduces down to $2n^2+1$ but the number of transitions shoots up to $n^4+2n^2$ (details in Table 3).

# 6 Conclusion

We presented a notion of $(i, j)$-predictability, an extension of predictability that specifies that there exists a time interval during which the fault occurrence is bound to happen in the system. This notion is very useful because it allows one to express different type of predictability, namely whether a fault can be predicted well in advance, whether the time of failure can be precisely predicted, or both.

There are several obvious extensions to these works, mainly regarding the expressive power of the modelling framework. We want to extend this work to timed systems [CG13], to probabilistic systems [NDY14], or to hybrid systems [BTO08]. Other works include the extension of the current work to decentralised predictors [TK12], the study of optimal observability for predictability akin to that of diagnosability [BLD08] or in combinaison with opacity constraints [CMPM14].

## Acknowledgments

NICTA is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program.

## References

[BLD08] L. Brandán Briones, A. Lazovik, and Ph. Dague. Optimal observability for diagnosability. In *Nineteenth International Workshop on Principles of Diagnosis (DX-08)*, pages 31–38, 2008.

[BM11] L. Brandán Briones and A. Madalinski. Bounded predictability for faulty discrete event systems. In *30th International Conference of the Chilean Computer Science Society (SCCC-11)*, 2011.

[BTO08] M. Bayoudh, L. Travé-Massuyès, and X. Olive. Coupling continuous and discrete event system techniques for hybrid system diagnosability analysis. In *Eighteenth European Conference on Artificial Intelligence (ECAI-08)*, 2008.

[CG13] F. Cassez and A. Grastien. Predictability of event occurrences in timed systems. In *Eleventh International Workshop on Formal Modeling and Analysis of Timed Systems (FORMATS-13)*, 2013.

[CL99] Ch. Cassandras and S. Lafortune. *Introduction to discrete event systems.* Kluwer Academic Publishers, 1999.

[CMPM14] S. Chédor, Ch. Morvan, S. Pinchinat, and H. Marchand. Diagnosis and opacity problems for infinite state systems modeled by recursive tile systems. *Journal of Discrete Event Dynamical Systems (JDEDS)*, pages 1–24, 2014.

[GL06] S. Genc and S. Lafortune. Predictability in discrete-event systems under partial observation. In *Sixth IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes (SafeProcess-06)*, 2006.

[GL09] S. Genc and S. Lafortune. Predictability of event occurrences in partially-observed discrete-event systems. *Automatica (Automatica)*, 45:301–311, 2009.

[JHCK01] S. Jiang, Z. Huang, V. Chandra, and R. Kumar. A polynomial algorithm for diagnosability of discrete-event systems. *IEEE Transactions on Automatic Control (TAC)*, 46(8):1318–1321, 2001.

[JMGL08] T. Jéron, H. Marchand, S. Genc, and S. Lafortune. Predictability of sequence patterns in discrete event systems. In *Seventeenth IFAC World Congress (WC-08)*, pages 537–543, 2008.

[JMPC06] T. Jéron, H. Marchand, S. Pinchinat, and M.-O. Cordier. Supervision patterns in discrete-event systems diagnosis. In *Seventeenth International Workshop on Principles of Diagnosis (DX-06)*, pages 117–124, 2006.

[NDY14] F. Nouioua, Ph. Dague, and L. Ye. Probabilistic analysis of predictability in discrete event systems. In *25th International Workshop on Principles of Diagnosis (DX-14)*, 2014.

[Rin07] J. Rintanen. Diagnosers and diagnosability of succinct transition systems. In *20th International Joint Conference on Artificial Intelligence (IJCAI-07)*, pages 538–544, 2007.

[SSL+95] M. Sampath, R. Sengupta, St. Lafortune, K. Sinnamohideen, and D. Teneketzis. Diagnosability of discrete-event systems. *IEEE Transactions on Automatic Control (TAC)*, 40(9):1555–1575, 1995.

[TK12] Sh. Takai and R. Kumar. Distributed failure prognosis of discrete event systems with bounded-delay communications. *IEEE Transactions on Automatic Control (TAC)*, 57(5):1259–1265, 2012.

[VTPS15] J. Vento, L. Travé-Massuyès, V. Puig, and R. Sarrate. An incremental hybrid system diagnoser automaton enhanced by discrenibility properties. *IEEE Transactions on Systems, Man, and Cybernetics (TSMC)*, 45(5):788–804, 2015.